Accepted for presentation at KR-MED 2008
draft material -- please do not cite

SNOMED CT: Browsing the Browsers

J Rogers[1], MD, O Bodenreider[2], MD,
[1]Technology Office, NHS Connecting for Health, Leeds UK
[2]National Library of Medicine, NIH, Bethesda USA
jeremy.rogers@nhs.net, olivier@nlm.nih.gov

SNOMED CT is a complex ontology; sophisticated browsers are required to make it understandable and useful. We identified 23 SNOMED CT browsers that have been developed, and inspected 17. We enumerate and provide test criteria for a 'master list' of 143 browsing features supported by at least one inspected browser; future work will determine which of these features are implemented by individual browsers. Only 5 features were common to all 17 browsers; 89 were found in less than one third of browsers. We recommend that a core set of browsing features be defined and harmonized across browsers, particularly for text-to-concept search operations.

INTRODUCTION

SNOMED CT is a biomedical ontology and an associated terminology[1.] Formerly owned by the College of American Pathologists, it has been managed since April 2007 by the International Health Terminology Standards Development Organisation (IHTSDO), a not-for-profit international standards body. As distributed, it is a large, complex and evolving knowledge artifact. Sophisticated browsers must make that complexity accessible and understandable, and suppress distracting or unwanted detail[2-3]. A number of different SNOMED CT browsers have been constructed since it was first published. Some have been evaluated for a variety of use cases, including coding of clinical data[4-8] and terminology evaluation and management[9].

In this paper, we report interim results of a systematic inspection of some of these browsers. We enumerate a superset of browsing features, outline the variability with which these features are implemented in individual browsers, and consider the possible consequences of non-standardized browsing of a standardized terminology.

MATERIALS

SNOMED CT

The core of a SNOMED CT release comprises three tables (sct_concepts, sct_descriptions and sct_relationships) collectively defining a compositional description logic ontology of the medical domain, and a lexicon of associated preferred or synonymous descriptions. The most recent international release (January 2008) contains 311,313 active concepts, 1,357,719 relationships between those concepts and 794,061 active descriptions.

Working deployments of SNOMED CT require additional or ancillary information linked to that core, usually provided by either the IHTSDO or a National Release Centre. Examples of such data include crossmaps to other clinical classifications (e.g. ICD-10), definitions of subsets of concepts and/or their descriptions for navigational or localization purposes, and a history of changes between successive releases. The January 2008 IHTSDO release therefore comprised 21 discrete table components in addition to the 3 defining the core ontology. The April 2008 UK National Release, which builds on the January 2008 IHTSDO release, comprised 122 separate tables.

In addition to this centrally provided additional content, it is also possible to link external data to the core or ancillary data sources. For example, crossmap target codes can be linked to their corresponding native rubrics or hierarchies.

SNOMED CT Browsers

The authors and their colleagues identified 23 different implementations of software[10-28] offering SNOMED CT browsing capability – either embedded in larger application environments or available as standalone browsers. 16 of these[10-23] were inspected as working software: CaTTS, CliniClue, CLIVE, EdBrowse, FDB Sphinx, HealthTerm, LexPlorer, Mycroft, NCI Terminology Browser, OntoBrowser, OpenKnoME, Protégé-OWL, SNOB, SnoFlake, the UMLS Rich Release Format (RRF) Browser and the Virginia Tech Browser. One additional feature was identified on a screen capture of the AxSys browser.

AxSys, CLIVE, FDB Sphinx, HealthTerm and LexPlorer require user privileges to access; OntoBrowser and EdBrowse are unsupported in-house prototypes. The remaining ten browsers are publicly available at zero cost. Both CliniClue and OpenKnoME require proprietary additional tooling to load SNOMED CT distribution files, although prebuilt CliniClue data is widely available. OpenKnoMe and OntoBrowser also require a proprietary terminology server.

The remaining 6 browsers not inspected[24-28] were: proprietary software from Informatics inc, Ocean Informatics and Visual Read; a demonstrator browser/encoder developed within the NHS Common

User Interface Project; Kermanog's CLAW product[17] based on SNOMED in ClaML (EN 14463) format; and Linköping University's browser. These were excluded for reasons of time or lack of access.

METHODS

Each browser was inspected by one author against an emerging catalog of all features exhibited so far by at least one previously inspected browser. Whenever the choice was given to us, browsers were inspected using content based on the July 31, 2007 international release of SNOMED CT. A subset of SNOMED CT content converted into OWL DL was used for Protégé-OWL inspection.

The goal of each successive inspection was primarily to identify novel features implemented in the inspected browsers, for inclusion in a cumulative master catalog. The feature catalogue was iteratively organized by an emerging set of themes, and this resulted in a progressive systematization of the inspection process itself, with each theme considered in detail by turn. This iterative systematisation aided the process of new feature identification.

Where possible, operational definitions of new features were specified (reproduced in Tables 1-3). Subsequent inspections progressed by browsing or searching the Test Case column entry, and comparing the displayed result with the Expected Result column. Although previously inspected browsers were subsequently re-inspected for newly discovered features, work is underway to confirm the validity and reproducibility of inspecting individual browsers against the feature catalog. Individual browser scores are therefore not presented here.

RESULTS

143 different browsing features were identified across 17 inspected browsers. 6 further features occurred to the authors during the inspection process as being potentially useful, but were not found in any inspected browser. The combined set of 149 features are presented in the accompanying tables, organised under the 8 major themes outlined below.

Our preliminary summary results, based on partially validated individual browser inspections, suggest most browser featuresets are an arbitrarily selected and small subset of all 149 features available. On average, individual browsers implement only 40 features (Range 21-107, StDev=13), but only 22 of the 149 features were found in more than two thirds of all browsers inspected, of which only 5 were implemented in **all** inspected browsers (Search by ConceptID or by Exact string, display of a ConceptID, its linkage to a Description, and the text of that Description). 89 features were found in less than a third of all browsers, but 70 of these are found in at least two browsers. Overall, these results suggests that most possible browsing features have been implemented independently by several SNOMED browser developers, but they have yet to become 'standard' across most browsers.

Core Data

A minimal requirement for a SNOMED CT browser is to give access to the data in the three core tables (concepts, relationships, descriptions). Table 1 lists the 22 fields from each of the three core tables that might be displayed by a browser.

Most browsers implement a concept-centric view of this core content, comprising one concept, its description(s), classification with respect to other concepts, and definition in terms of other concepts. This represents the minimum set of features required for the coding of clinical data and basic navigation.

Some fields (e.g. ConceptStatus) appear in the source release data as coded numeric values whose interpretation is given only in SNOMED release documentation; most browser implementations display only the human readable interpretation of these codes and not also (or only) the numeric values as actually distributed.

Despite their 'core' nature, however, only three of the 22 related features were displayed by all browsers inspected: the Concept ID, a link to (at least one) description for a concept, and display of the text of linked descriptions. Description status and Initial Capital Status, Relationship ID and Refinability were each visible in only two or three browsers.

Non-Core: Ancillary, 3rd Party and Derived Data

Advanced navigation and terminology maintenance work may require either additional data outside the core tables, or 'derived' views of the core data itself such as 'reverse' historical relationships (showing which inactive concepts point at the current browser focus concept as their replacement). Table 1 lists the 'derived' views found across the inspected browsers.

A complete set of SNOMED core and ancillary linked data is large and complex. Further, it changes with each biannual release. To reflect this configuration and versioning complexity, some browsers report exactly which versions of which release components are loaded, alert users when they are browsing non-current data, and support concurrent browsing of multiple release versions for direct discovery or comparison of changed content.

We found display of non-core data, and data from more than one release, to be the exception rather than the rule. Pointers from inactive concepts to their active replacement, and the set of concepts using the browser focus concept in their definition, are accessible in less than half of all browsers; all other ancillary, 3rd party or derived data browsing functions are present in less than one third of all browsers and usually only in two or three.

Visualisation and Navigation

Following from consideration of what data a browser displays is *how* it displays it. Additionally, the navigability of this data must be considered. Table 2 lists the visualization and navigation features encountered in the inspected browsers.

Most browsers implement some form of graphical tree browser, displaying the browser focus concept in the context of SNOMED's multiaxial subsumption hierarchy. Some off-the-shelf tree controls, however, are unsuitable for displaying trees with very many levels and very many siblings at the same level, such as SNOMED CTs subsumption hierarchy. Those showing the hierarchy always exploded from the root node downward (e.g. the NCI Terminology Browser and Protégé) are particularly unwieldy; those that do not detect very large sibling sets before attempting to display them can lead to very long refresh times.

Other visualization features observed include: sorting and grouping of components within concept definition or synonym sets, diacritic and superscript rendering, and typographic or colour coding of text.

Most browsers employ web browsing paradigms for navigation, with use of hyperlinks to refocus the browser on arbitrary concepts, as well as back/forward navigation. Bookmarked 'favourites', or a 'home' concept, however, were rarely observed.

Usability and Interoperability

The overall experience of working with a browser is influenced by a range of more generic user interface features, listed inTable 2. These include: the ability to transiently or persistently configure a custom view on the wealth of SNOMED related information, e.g., to occupy less of the desktop real estate; copy-and-paste or drag-and-drop of selected information either within the browser environment or into external applications, and the availability of an API allowing browser interface components to be instantiated and controlled by 3rd party software (a functionality distinct from the notion of a terminology services API per se).

Searching

Table 3 lists the range of features observed by which SNOMED CT is searched against a user-entered text string in order to identify candidate SNOMED ConceptIDs as possible entry points for subsequent visualization and navigation. These different search features observed may be further analysed into:

- lexical expansion of the original user search string in order to increase recall
- semantic or metadata filtering of the set of candidate concepts returned by a query, in order to increase precision
- collation and sorting of filtered results, so that the user may find (or be certain of **not** finding) the required concept

In general, SNOMED CT searching functionality in most browsers is impoverished and idiosyncratic. Although 37 different query expansion, filtering and collation features were observed across all browsers, thirteen of the browsers implemented less than 10 of them - and rarely the same set. 27 searching features were implemented in less than a third of all browsers inspected, of which 5 were unique to one browser.

Browsers differ in which features are on by default, which must be explicitly specified, and which can be, or by default are, combined in Boolean combinations. Not all strip trailing spaces; some default to an exact string match whilst others assume wildcarding unless specifically overridden. Where a search expression contains multiple words or tokens, few browsers support complex query logics such as requiring some tokens to be present and others not.

To demonstrate the effect of these differences, all browsers were used in their default configuration to search against the same string: 'ear catheter'. Six browsers found no matches. A further six found only 72683003 Removal of catheter from middle ear, and its two descendants. SNOB returned eleven matches, including 72683003 but also 232199004 Inflation of Eustachian tube using balloon. The latter has no directly associated descriptions containing either 'ear' or 'catheter' but instead is returned because it has at least one ancestor with at least one description matching 'ear', and a separate ancestor with a description matching 'catheter'. The UMLS RRF Browser returned sixty-six matches.

Postcoordination and Miscellaneous

Unlike traditional clinical terminologies, SNOMED CT can be 'postcoordinated' - dynamically extended by anybody, subject to certain ontological rules. Most trivially, this manifests as the option to qualify anatomical sites by a *Laterality* attribute and *Sidedness* value. Exposing SNOMED CT only as a static corpus significantly diminishes its expressivity. Further, a large part of the content – e.g. all Qualifier, and Linkage Concepts - is easily misunderstood outside the context of postcoordination.

The rules governing postcoordination are complex but compliance with them is a prerequisite for dynamic classification of the expressions so built. A dedicated postcoordinated expression building and validating interface is therefore highly desirable, but we found only five browsers that implement one. Three of these additionally implement some limited part of the rules and conventions. However, although compliance with the rules has limited value outside the context of dynamic classification, no browser inspected currently provides that function.

SNOMED CT contains many content errors and omissions. Empowering end users to log and report content errors offers a 'social computing' route to expand SNOMED CT's quality assurance capacity. However, only one inspected browser directly integrates content bug logging and reporting.

DISCUSSION

Accessing data vs. browsing. In seeking to review 'browser' technologies, we excluded command line or other direct SQL interfaces on the data tables. Although most browsers hide the raw data tables from the user, at least one explicitly provides a route to it. Whether 'display' of data by this route should pass or fail our core data theme tests is debatable.

Configurability. A minority of the features identified are orthogonal or graded values of one property. For example, whether a given hierarchy browser sorts sibling concepts randomly, alphabetically by description, or numerically by ConceptID are orthogonal values of a 'sibling sort' function. Although in theory it is possible to imagine a browser configurable to any one of the three, individual hierarchy display instances can only implement one at a point in time. In practice, all inspected browsers implement only one of these options throughout.

Operational test criteria. Differences between the browsers, particularly their default treatment of search strings, confounded attempts to specify tests that would work equally across all of them. Many of the tests specified in Tables 1-3 must be interpreted to take account of issues such as whether exact or wildcard string matching is assumed.

Absence of standard search features. The observed differences in text-to-concept search implementations have a striking effect on browsing experience. Further work to characterize this phenomenon is required.

Future work. We are currently validating the testing of specific browsers against the catalog of features. The quantitative results reported here are preliminary but confirm the authors' original motivation for the experiment: currently available SNOMED CT browsers are very different and often suboptimal.

We do not propose that all SNOMED CT browsers must always implement all the features we identify; further research is required to determine which features are required for specific use cases, but the prior existence of a master feature catalog such as we present here is a prerequisite for that research. Many of the features seem likely to be common across use cases, particularly text-to-concept search operations. We recommend that a core set of searching and browsing features be defined and harmonized across tools, so that a standard terminology is not transformed into multiple different objects by virtue of idiosyncratic and limited browsing experiences.

Acknowledgments

This research was supported by the Intramural Research Program of the National Institutes of Health (NIH), National Library of Medicine (NLM) and by NHS Connecting for Health. Thanks to Drs Malcolm Duncan and Christopher Wroe for their assistance.

References

1. SNOMED CT. IHTSDO, Copenhagen 2007 www.ihtsdo.org
2. Tuttle MS, Cole WG, Sheretz DD, Nelson SJ. Navigating to knowledge. Methods Inf Med. 1995 Mar;34(1-2):214-31
3. Patel VL, Kushniruk AW. Understanding, navigating and communicating knowledge: issues and challenges. Methods Inf Med. 1998 Nov;37(4-5):460-70.
4. Windle J, Van-Milligan G, Duffy S et al. Web-based physician order entry: an open source solution with broad physician involvement. AMIA Annu Symp Proc. 2003;:724-7.
5. Elkin PL, Brown SH, Husser CS et al. Evaluation of the content coverage of SNOMED CT: ability of SNOMED clinical terms to represent clinical problem lists. Mayo Clin Proc. 2006 Jun;81(6):741-8.
6. Sundvall, E., Nystrom, M., Petersson, H., Ahlfeldt, H. Interactive visualization and navigation of complex terminology systems, exemplified by SNOMED CT. Stud Health Technol Inform. 2006;124:851-6.
7. Chiang MF, Hwang JC, Yu AC et al. Reliability of SNOMED-CT coding by three physicians using two terminology browsers. AMIA Annu Symp Proc. 2006;:131-5.
8. Richesson R, Syed A, Guillette H et al. A web-based SNOMED CT browser: distributed and real-time use of SNOMED CT during the clinical research process. Medinfo. 2007;12(Pt 1):631-5.
9. Cornet R, de Keizer NF, Abu-Hanna A. A framework for characterizing terminological systems. Methods Inf Med. 2006;45(3):253-66.
10. CaTTS (browsed Dec 21st 2007) www.jdet.com/
11. CliniClue (build 2006.2.30) www.cliniclue.com
12. CLIVE (UK NHS in-house terminology authoring tool)
13. HealthTerm (v 4.3.2 browsed Dec 21st 2007)
14. HLi LExPlorer (v 4.4.1P build 48 browsed Dec 21st 2007 – Athens account required) www.snomed.cfh.nhs.uk/lexplorer/
15. Mycroft (v. 2.1.0.2) www.apelon.com/
16. NCI Terminology Browser (browsed Dec 21st 2007) nciterms.nci.nih.gov/NCIBrowser/
17. OpenKnoME 5.4d and ClaW Workbench www.opengalen.org/sources/software.html

18. Protégé (v4.0 build 59) protege.stanford.edu
19. SNOB (v1.64) snob.eggbird.eu
20. SnoFlake (v 2.0 browsed Dec 21st 2007) snomed.dataline.co.uk/
21. UMLS Rich Release Format Browser (2007AC) www.nlm.nih.gov/research/umls/
22. Virginia Tech Browser (browsed Dec 21st 2007) terminology.vetmed.vt.edu/SCT/menu.cfm
23. AxSys Browser (browsed Jan 5th 2008) www.axsys.co.uk/excelicare/eprclinicalcoding.htm
24. First DataBank www.firstdatabank.com/
25. Informatics inc www.informatics.com/
26. Ocean Informatics oceaninformatics.biz
27. Visual Read www.visualread.com/
28. NHS Common User Interface nww.cui.nhs.uk/

FEATURE	TEST CASE	TEST DETAILS
CORE SNOMED CT DATA		
Basic Concept Table Info		
Fully Specified Name	105531004	Browser indicates clearly that the FSN for this concept is: Housing unsatisfactory (finding)
ConceptID	105531004	When browsing this concept, the conceptID 105531004 is clearly displayed as being the focus of the browser
ConceptStatus (original numerical value)	174463006	Browser indicates that the ConceptStatus for this concept is: 6
ConceptStatus (human readable text - current, ambiguous etc)	174463006	Browser indicates that the ConceptStatus for this concept is: Limited
IsPrimitive	105531004	Browser indicates that the IsPrimitive value for this concept is: True (numerical vaue = 1)
CTV3ID	105531004	Browser indicates that the CTV3ID for this concept is: XaBzq
SNOMEDRT ID	105531004	Browser indicates that the SNOMEDID for this concept is: S-31232
Basic Description Table Info		
DescriptionID	37810007	Browser indicates that the preferred Term 'Myeloid leukemia' has DescriptionID = 486867011
ConceptID	37810007	37 current descriptions linked to this ConceptID can be displayed; it is clear which ConceptID they belong to
DescriptionStatus (original numeric value)	37810007	Browser indicates that the term 'Myelocytic leukemia, NOS' is now DescriptionStatus=1 for this concept
DescriptionStatus (current, erroneous, retired etc)	37810007	Browser indicates that the term 'Myelocytic leukemia, NOS' is now 'retired' (DescriptionStatus=1) for this concept
DescriptionType (preferred, synonym, FSN)	37810007	Browser clearly identifies which descriptions are preferred term, which synonyms and which the Fully Specified Name
Term	369881000	Any human readable terms associated with a given concept are displayed IN FULL and without truncation
InitialCapitalStatus Flag Value Displayed	100000000	Browser indicates that the FSN for this concept has INITIALCAPITALSTATUS value: TRUE (1)
Language (en, en-gb, en-us etc)	37810007	Browser displays TWO alternate preferred terms, AND indicates that : 486867011 Myeloid leukaemia has LANGUAGECODE 'en-gb'
Basic Relationship Table Info		
RelationshipID	235583009	Browser indicates that the defining relationship (Method = Incision) has RelationshipID=1795591025
ConceptID1 (or FSN/Preferred term)	235583009	Where the browser displays a relationship, the concept modified by that relationship is clearly identifiable and its term and/or ID are displayed
RelationshipType (or FSN/Preferred term)	235583009	Where the browser displays a relationship, the attribute involved in that relationship is clearly identifiable; its term and/or ID are displayed
ConceptID2 (or FSN/Preferred term)	235583009	Where the browser displays a relationship, the concept that is the value of the relationship is clearly identifiable; its term and/or ID are displayed
CharacteristicType (defining vs optional vs additional)	86299006	Browser indicates that the relationship (Occurrence = Congenital) is defining, while (Severity = Severities) is an optional qualifier
Refinability	307244005	Browser indicates that the optional qualifier relationship (UsingDevice = Balloon dilatation catheter) has refinability status 'not refinable'
RoleGroup	86299006	Browser shows concept has five role groups plus role group zero.
ANCILLARY, 3rd PARTY AND DERIVABLE DATA		
Ancillary Table Info		
Crossmaps	72683003	Browser shows some possible crossmaps from at least one external classification e.g. to D20.3 in OPCS 4, 4.3 and 4.4
Subset membership	160573003	Browser shows this concept to be a member of UK Alcohol subset (NB IHTSDO release only, use 108928007 in US Proprietary Drugs Subset)
Concept history (when added etc)	412060000	Browser indicates this concept was first added in the 20040731 release of SNOMED CT
Namespace/extension identification	4702411000001104	Browser indicates this concept belongs to the UK Drug extension namespace
3rd Party Info		
Original SNOMED ID Preferred Term	3419005	Browser displays the native SNOMED RT preferred term for 'DE-11720', (=SNOMEDID) (should be 'Faucial diphtheria')
Original CTV3 Preferred Term	111487009	Browser displays the the native CTV3 preferred term for 'E2749 (=CTV3ID) (should be 'Nightmares')
Original external scheme terms	3419005	If the browser displays ICD10:A36.0 as a crossmap, it ALSO indicates that the rubric for A36.0 in ICD10 is 'Pharyngeal diphtheria'
Hyperlink to Xmapped scheme browsers		Where a mapping to an external scheme is shown, this is hyperlinked to a browser on that scheme
Allowable Concept Model Relationships	3419005	Browser indicates that this concept can ALLOWABLY be qualified by (Occurrence = Periods of Life), amongst other possibilities
Dynamic flagging of modelled relationships outside concept model	403600002	Browser indicates that the stated defining relationship (Has definitional manifestation = Formication) does not comply with SCT concept model
Derivable Table Info		
All concepts that use X in their definition	129123002	Browser should indiate link to 4 other concepts: 52814001, 112890006, 21279007 and 75667007
Forward history relationships	7222009	Browser shows one REPLACED_BY relation to 6081001 Deformity (NB test only possible using Jan 2005 release or later)
Reverse history relationships	72683003	Browser shows 3 MAYBE and 1 SAMEAS relations from inactive concepts
Content Metrics		Browser can display summary stats (total concepts/relationships/descriptions, of which active/inactive etc)
Reverse cross maps	V60.9	Browser returns 4 different ConceptIDs that are mapped directly to this code in ICD-9-CM
Code conversion (single SCTID -> nearest target scheme map)	275875002	Concept has no direct ICD-10 map; nearest ancestor with one is [239958005 Painful arc syndrome] which is mapped to M75.1
All members of one subset	2431000000138	Browser displays all members of UK smoking subset: this has 6 top level members (and others in descent of those 6)
Short normal form	182555002	71388002|Procedure|:{363699004|Direct device|=63995005|Bandage|,260686004|Method|=129425003|Application - action|}
Long normal form	286572006	286572006|Activation of implant|:363699004|Direct device|=40388003|Implant|
Dynamic inheritance of relationships	320630002	Should show [HasDoseForm =Pressurised inhalation] inherited from parent; inhalation powder isn't a subtype of pressurised inhalation
Similarly indexed concepts	37810007	Browser allows you to select 2 synonyms - granulocytic, and eosinophilic leukaemia - an autoconstructs a new search expression on their union
Other concepts with exact same description(s)	26239002	Preferred term 'football' is shared as description ONLY on 413489002, 413492003, 413494002 and 88289009
SNOMED Release Versioning		
Names of all source files actually loaded		Browser includes option to display full listing of all original source files loaded
Currency of content (whether release is outdated)		Browser prominently display some indication of the content version being browsed, and whether it is 'current'
Simultaneous browsing of any set of different SCT releases		Browser supports simultaneous AND mutually interactive browsing of more than one version of SNOMED content

Table 1 Core and Additional SNOMED CT table browsing features

FEATURE	TEST CASE	TEST DETAILS
VISUALISATION		
Hierarchy And Other Content Visualisation		
Indefinitely Expandable/Collapsible Ancestor Tree	107944001	Browser hierarchy can be indefinitely expanded looking upward to reveal all ancestors of focus concept
Indefinitely Expandable/Collapsible Descendant Tree	107944001	Browser hierarchy can be indefinitely expanded looking downward to reveal all descendants of focus concept
Indefinitely Expandable/Collapsible Bidirectional Polyhierarchy Browser	107944001	Single hierarchy browser instance supports both of above
Bidirectional transitive closure only browser	107944001	Single hierarchy browser instance can display all descendants, and all ancestors, but WITHOUT also displaying all siblings of all ancestors
Descendant Hierarchy does not have to start from SNOMEDCTConcept	107944001	Single tree view instance can display hierarchy of ALL descendants of a concept without also seeing all its ancestors up to SNOMEDCT Concept
Large result warning and bail-out	363662004 / 102272007	DuplicateConcept has 58k children; other test concept has 2530. Browser offers bail out if it is going to take a a long time to retrieve and display.
Flag where same concept currently displayed >1 time same browser	363687006	Expand 'Arthroscopic procedure' and 'Endoscopic Biopsy': 'Arthroscopic synovial biopsy' appears twice
Alphasorting of synonyms	37810007 / 79962008	Browser lists displayed synonyms for a concept in alphabetical order
Alphasorting of siblings	102272007	Browser lists the sibling descendants of 102272007 in alphabetical order (NB accept upper / lower case interleaved, or all upper first then all lower)
Collation of relationships on same attribute within a role group	394878001	Browser groups all relationships of same name together in the displayed list of relationships
Relationships are not sorted (+/- collated) numerically by attribute conceptID	394878001	Score 'Y' if attributes are NOT displayed in the following sort order: finding site < interprets < finding method < finding informer
Relationships are alphasorted (+/- collated) by role name	394878001	Score 'Y' if attributes ARE displayed in the following sort order: finding informer < finding method < finding site < interprets
Sets of same relationships can't be sorted numerically by conceptID of value	394878001	Score 'Y' if values for [interprets] relations ARE NOT shown in order: Resp effort < Gen structure of thorax < Method of breathing < Resp function
Alphasorting within sets of same relationship by value name	394878001 / 125852008	Score 'Y' if values for [interprets] relations ARE displayed in order: Gen structure of thorax < Method of breathing < Resp effort < Resp function
Conceptual grouping or filtering of relationships by relationship supertype	239946005	Browser lists [After] and [Due to] relationships together in Role Group 0, because they are subtypes of [AssociatedWith]
Rendering of superscript/subscript descriptors	65527003	Browser displays Preferred Term as: 190m (in superscript) 1 (in subscript) Iridium
Rendering of diacritics	80734006 / 13445001	Browser displays sjögren not sjÃ¶gren / accented e's
Colour and typographic coding of hierarchy nodes by status etc		Use is made of colour AND typographics (italics, bold etc) to encode data such as primitive status, limited status, subset membership etc
Snoflake / Cloud / Other GUI views		Collections of concepts can be displayed using a graphical paradigm OTHER than a tree view widget
Hierarchies on non IS-A relationships	244355000	Hierachy can be configured to show IS_PART_OF hierarchy (ie 244355000 under 181286006 under 362008007 under 302509004 etc)
Dynamic Display Filtering		
By concept status	174461008	Browser allows you to dynamically suppress and reveal display of 174464000 and 174463006 as limited status children
By description status	105531004	Browser allows you to suppress the non-current descriptions
By description type	44054006	Browser allows you to dynamically suppress and reveal display of synonyms or preferred terms or fully specified names or any combination
By description language	196623008	Browser allows you to dynamically suppress and reveal the en-gb descriptions
NAVIGATION		
Hierarchy Navigation		
Browse history (back, forward buttons)		After clicking through a series of hyperlinks until browsing a different concept, BACK and FORWARD buttons can step through the browse trail
Refocus browser on any displayed concept (hyperlinked browsing)	3419005	Hyperlink on [5851001 Corynebacterium diphtheriae] in definition spawns another browser hierarchy (or refocusses current) on 5851001
Refocus browser on 'Home' concept		The entire interface can be quickly refocussed on SNOMEDCT, or some other user defined 'browser base' concept
Persistent user-defined favourites list		Users can declare and manage their own list of favourite browser base concepts, and these are persistent between sessions
USABILITY & INTEROPERATION		
User Interface Usability and Interoperability		
Internal Drag-and-drop		Browser elements can be refocussed by dragging and dropping a concept from another element
External Drag-and-drop (e.g. to MS Word, Outlook)		Human-readable concept representations (e.g. CG notation) can be dragged and dropped into external applications
Scroll-wheel enabled whenever any long list visible (e.g. tree view)	128303001	Long list of descendant concepts can be scrolled up and down. E.g. in the tree view, using a mouse wheel (if fitted)
Maximum single session hierarchy browsers		No. of hierarchy instances that can be spawned using controls within the UI (not by starting a new client)
Hierarchy browser expansion state memory	128303001	Expand several layers from Excision, then close at level of Excision and reopen; hierarchy below Excision is pre-expanded
Concept information widget state memory		Where a component is configured to display a subset of core and ancillary info, configuration is persistent across browsed objects (e.g. if synonym display off by default, stays on indefinitely if switched on by user)
Adjustable font size		Whether size of font for hierarchy and other text is user configurable
Browser component slaving on same or different databases		Browser component B can be slaved so when Browser component A refocusses on conceptX, so does B
Subsumption hierachy only view		The browsing user interface can be reduced to displaying ONLY one subsumption hierarchy, but the full environment may be recovered
Plug-in user extensible architecture / API		Users can integrate their own tools and search enhancements into the existing UI
Single Application Window Interface		All browser elements visible at once within a single Window object; no element can be minimised
MDI Container Interface		All browser elements exist within a single windowed object. They can be minimised, but not to the task bar or dragged to a virtual desktop
Multiple spawnable window interface		Browser elements exist as independent windowed objects. They can be independently minimised to the task bar or dragged to virtual desktops
Independently scalable windows		The X-Y dimensions of all major browser elements can be set independently of each other
User configurable suppression of IDs in tree view browser		Concept IDs can be displayed or not displayed in the hierarchy browser at user's discretion
Copy hierarchy as text		Any displayed subsumption hierarchy can be copied EXACTLY to clipboard as tab-indented ASCII text, preserving precisely which nodes are or are not expanded
Copy hierarchy as graphic		Any displayed hierarchy can be copied to clipboard as a pre-cropped screencap (no 3rd party screengrab required)
Copy hierarchy as XML / HTML / RTF		Any displayed hierarchy can be copied to clipboard as formatted text (XML, HTML, RTF)

Table 2: Visualisation, Navigation and Interoperation browsing features

FEATURE	TEST CASE	TEST DETAILS
SEARCHING		
Text-to-concept search techniques and filtering		
Search by ConceptID	105531004	Concepts can be searched for by direct entry of their conceptID, as well as by lexical string match against natural language words or phrases
Has (and/or does not have) concept status	Motor	Search can be filtered to return ONLY those concepts with status=6 (limited) and NOT those with other status values also
Has (and/or does not have) concept supercategory	Motor	Search can be filtered to ONLY return 2 concepts that are in [Special Concept] supercategory and NOT also those in other categories
Has (and/or does not have) ancestor = conceptID	Motor	Search can be filtered to show ONLY 1 concept that is a subtype of '415577004 Sport' and NOT also other concepts
Has (and/or does not have) relationship attX with %Y	endoscopic	AND NOT (procedureDevice=endoscope) returns 22 procedures with 'endoscopic' in description, but no modelled endoscope (7 if case sensitive)
Is (or is not) used in relationship attX from %Y	valve NOT PartOf anat.	(Jan 2008 data) should return 280646001 Valve of nasolacrimal duct, 110551008 Aortic and mitral valves, CS, plus 2 others, but not e.g. mitral valve
By namespace or originating authoring centre	Gelfoam	Search can be filtered to return only the 2 concepts that are in the 1000000 namespace (UK Extension), or one in the 1000002 namespace (USDRG)
By description status	Phytohemagglutinin	Search restricted to retired descriptions (status=1) returns 10669005 ONLY. (Unrestricted search also returns 252350000 and 36658009)
By description type	insane	Search restricted to preferred term returns zero results (Unrestricted search returns 6 results of which 3 are current)
By description language	tranquilizer in 'en-US'	(Jan 2008 data) returns 36 concepts without filtering to restrict results to en-US matches, and 7 with that filter applied
Is (not) member of subset	lamisil in Subset:109034	(Jan 2008 UK data) returns 28 active concepts if filtered to include only those in US Proprietary Drug Subset (40 with no filter)
Exact string match (including on multitoken strings)	[feeding education]	Only 4 CORE concepts are returned on exact match; 6 will be returned by a search on 'breast' AND 'fed' (7 if case insensitive search)
Partial string match (wildcarded)	tetral*	Browser returns at least 86299006 Tetralogy of Fallot (may also return other matches)
Force exclude/include terms that match token	tetral* -shunt	Browser does NOT return 12363009 Complete repair of tetralogy of Fallot with closure of previous shunt
Force exclude/include matches by frequency of token in index		
Phonetic query expansion (Metaphone, SoundEx etc)	epididimiss	Browser either automatically substitutes correct spelling 'epididymis' or prompts user for likely substitution
Stemming and part of speech lexical substitute query expansion	toys	Search on 'toys' will return 97 results if the plural is being normalised to 'toy', else only one result
Strip trailing/leading spaces & non-alpha characters on conceptID	3419005	Any concept ID followed by white space or a comma
Word concatenation query expansion	wheel chair	Browser returns 225612007 (under clinical findings) as well as other results
Colloquial term or external term substitution query expansion	fetal	Browser returns [401091000 Cold agglutinins level : foetal cells (procedure)] as first level concept, amongst others
Stopword list (query contraction by blacklist)	between / specified	No results (as opposed to the expected hundreds of results) are returned to single word using: which, between, under, specified etc)
Token separators other than WS in descriptions	DOC	Exact word match returns [131459009 Cham-Doc cattle] and [1336006 11-Deoxycorticosterone]
Force case sensitive/insensitive searching	ACE vs ace	Case sensitive search on ACE will not NOT also return Ace bandage (16568017)
Handling of diacritics	Lowchen / Ménière	[Lowchen] returns [132607008 Löwchen dog] despite no unaccented match; [Ménière] returns 252546006 despite no accented match
Indexation by isA relationship to concepts with matching descriptions	ear catheter	Browser returns [232199004 Inflation of Eustachian tube using balloon] as one of the search results
Indexation by non-isA relationship to concepts with matching descriptions	contagious adrenal	Browser returns [11244009 Polyglandular autoimmune syndrome, type 1] as one of the search results
Boolean combinations		Complex searches can be specified as boolean combinations of all above
Full regular expression syntax for search expression	%M_ni_re% or equiv	Browser returns 60 or so results, including 252546006, matched on both Meniere and Ménière variants, plus matches on 'Minipress'
Type ahead autocompletion / suggestion		Search box, or result box, dynamically recomputes possible matches as you type
Grouping/nesting of results by supercategory / subsumption	prosthesis*	430 results are spread across 9 SNOMED supercategories: are all results for one supercategory grouped together?
Returns matching descriptionID rows	ear AND catheter	Returns two instances of 72683003 - one for preferred term, one for FSN
Returns non-redundant set of concepts with matching descriptions	ear AND catheter	Returns one instance of 72683003
Search match ranking by frequency of clinical enduser usage		Browser displays search results with those most frequently used by a group of users at the top
Search match result sorting by character length		Browser displays search results ordered according to the number of characters in the displayed strings
Search match accuracy scoring and ranking/sorting by score		Browser displays search results ordered according to some metric for the accuracy of the match to the original search string
Result truncated, or paged, by user defined limit (e.g. to max 500 results)		The user can set a threshold maximum number of results to be returned, or displayed at the start, to avoid very large result sets
Result throttling by user defined match accuracy score threshold		Where the browser computes a match accuracy value on each match, the user can supress all results below a threshold score
No. Results to 'ear catheter' (active concepts only) with default search	ear catheter	Result = all concepts returned PLUS all their descendants on July 2007 content
POSTCOORDINATION		
Postcoordinated Expressions		
Expression builder UI		The browser includes a UI element for building novel modelled expressions
Allowable choices offered predictively	3419005	Browser notifies you that you can allowably refine this concept by (Occurrence = Periods of Life)
Refinable choices offered predictively	3419005	Browser notifies you that you can refine the existing defining relationship (FindingSite=Fauces structure)
Refinability flag aware	173345004	Browser forces you to specify a subtype of SurgicalAccessValues, but won't allow you to specify a subtype of Device for UsingDevice
Allowable/Refinable/Optional candidate forecasting	3419005	Browser doesn't offer you the possibility to refine the existing defining relationship (AssocMorph=Pseudomembrane) (pointless; no descendants)
Nested expression building	3419005	Browser allows you to refine the existing FindingSite to 21294006|Palatine arch structure, AND then lateralise it
Role group aware	3419005	Browser places any refinement of the defining relationship (FindingSite=Fauces structure) in Role Group 1
Postcoordinated expression validation against concept model		Externally created post-coordinated expressions (e.g in CG notation) can be validated against concept model
Dynamic classifier		Externally or internally post-coordinated expressions can be classified against loaded content
MISCELLANEOUS		
Content error reporting		
Internal content error reporting		The browser includes a tool to persistently annotate concepts or groups of concepts with comments or errors etc
Internal content error management		Concept annotations can be edited, grouped, organised, deleted, imported and exported
External content error reporting		The browser includes a tool to share content error reports or comments externally to the browser/user instance
External content error syndication		The browser can automatically syndicate (push and pull) error reports and comments to and from other users
Other User Interface Features		
Subset editor and exporter		The browser includes a UI element for creating, or editing, subsets such as may be used to filter browsing behaviour
'TQL' Parser		The browser includes a comprehensive command line Terminology Query Language compiler/parser
Standalone (does not require internet connection)		SNOMED content is resident on the users computer
Bail-out to view on raw tables		Browser displays the raw tabular release ASCII files
Distributed independent of data (user can rebuild)		Users can rebuild customised SNOMED content locally (or on a remote server)

Table 3: Searching, Postcoordination and Miscellaneous browsing features